



Train Station Hazard Detection

Ein KI-basiertes Sicherheitssystem zur Gefahrenerkennung an Bahnhöfen

Motivation

Kind in Frankfurt vor einfahrenden ICE gestoßen – Junge (8) tot



2018: 2401 Verletzte und 185 Tote im Schienenverkehr

Für jeden ICE ist "Personen im Gleis" mindestens einmal pro Woche Grund für Verspätungen. Insgesamt kommen bis zu 200 Verspätungsminuten zusammen. Eine Verspätungsminute im Fernverkehr kostet die DB 100€. 2019 gab es insgesamt 289 ICE. Damit ergeben sich Kosten von bis zu **300 Mio. € pro Jahr** durch Verspätungen wegen Personen im Gleis.

Angenommen 5% der Vorfälle geschehen an Bahnhöfen und von diesen Vorfällen könnten wir annähernd 90% durch Realzeit-Warnungen verhindert werden.

Diese Warnungen würden damit bis zu **13,5 Mio. € pro Jahr** an Verspätungskosten einsparen.

Unsere Herausforderung:

Automatisiertes Realzeit-Warnsystem erstellen, um Unfälle und Tote zu vermeiden, sowie Kosten durch Verspätungen um bis zu 13,5 Mio. € zu reduzieren

Unsere Lösung

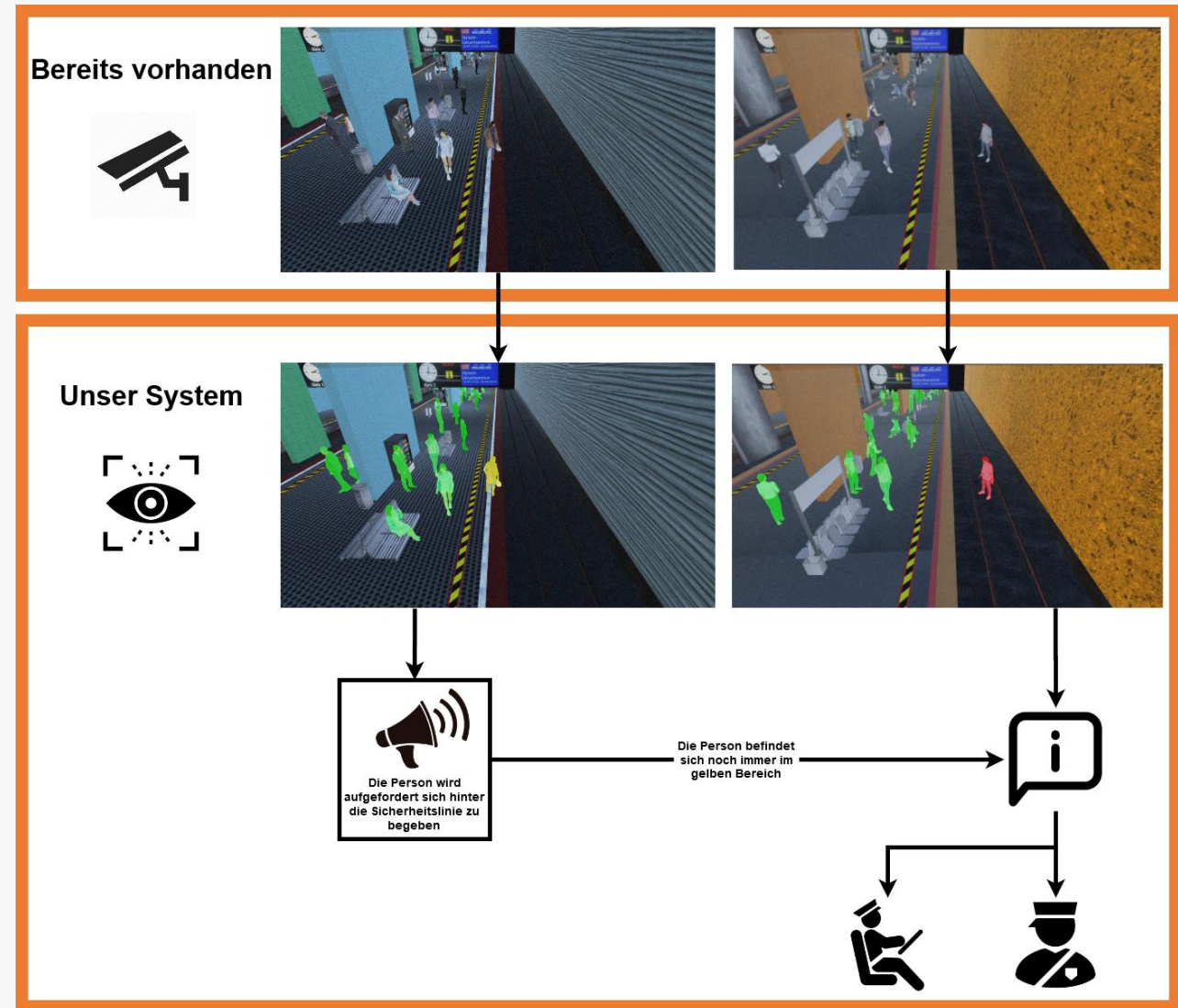
Für die automatisierte Gefahrenerkennung nutzt unser KI-basiertes Sicherheitssystem die bereits vorhandene Kamera-Infrastruktur an Bahnhöfen und ermöglicht somit eine kosteneffiziente Umsetzung der Lösung. Dies wiederum ermöglicht Gefahrensituationen frühzeitig zu erkennen und entsprechende Präventionsmaßnahmen zur Unfallverhinderung einzuleiten. Dabei ist die Aufgabe des Systems die Menschen auf den Kamerabildern zu erkennen und in 3 Sicherheitskategorien zu unterteilen:

- Grün (keine Gefahr)
- Gelb (potentielle Gefahr)
- Rot (absolute Gefahr)

Zur Unterscheidung zwischen grünen und gelben Personen wird an jedem Bahnhof eine Sicherheitslinie angebracht, welche sich genau 1,5m von der Bahnsteigkante entfernt befindet.

Sobald eine Person die Sicherheitslinie überschreitet, ohne dass ein Zug eingefahren ist, wird diese Person durch das Kamera-System erkannt und mit einer Durchsage aufgefordert diesen Bereich zu verlassen. Falls die Person den Bereich nicht verlässt, wird sie als potentielle Gefahr erkannt und sowohl der Lokführer als auch ein Sicherheitsbeamter informiert. Durch diesen Ansatz mindern wir das Risiko durch Stoßen und erhöhen die allgemeine Sicherheit an Bahnhöfen.

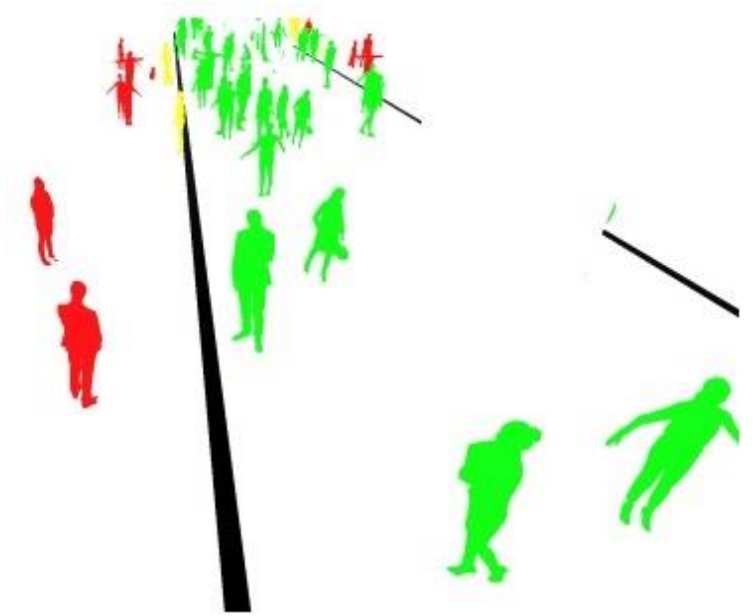
Sollte eine Person das Gleisbett betreten, wird sie automatisch als absolute Gefahr erkannt. Sowohl der Lokführer als auch ein Sicherheitsbeamter werden in diesem Fall informiert und können somit ggf. Maßnahmen zur Unfallprävention einleiten.



Erster Prototyp – Simulation

Um unser System auf alle möglichen Gefahrensituationen zu trainieren und zu testen, haben wir eine 3D Simulation für U-Bahn-Stationen entwickelt. Die Simulation besteht aus 10 Grundbahnhofsf lächen, welche alle möglichen Kombinationen von Schiene, Wand und maximal 2 Bahnsteigen abdecken kann. Bei der Erzeugung einer Bahnhofsszene wird aus einem Auswahlpool von über 100 unterschiedlichen Texturen, sowie mehreren Mülleimern, Sitzmöglichkeiten, Fahrstühlen, Roll/-Treppen, Anzeigetafeln und Snackautomaten ausgewählt, um einen realistischen Bahnhof zu kreieren. Anschließend werden Menschen anhand unserer auf echten Daten basierenden Verteilung am Bahnhof platziert. Zudem wird an jeder Bahnsteigkante eine Sicherheitslinie in 1,5m Abstand zur Bahnsteigkante angebracht, welche Teil unseres Sicherheitssystems ist.





Erster Prototyp – Train Station Segmentation

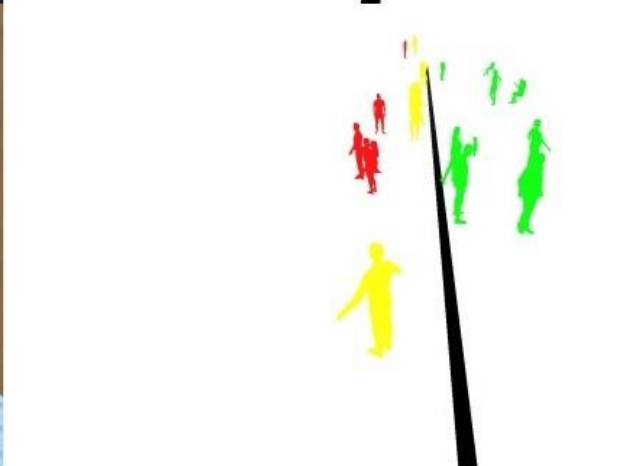
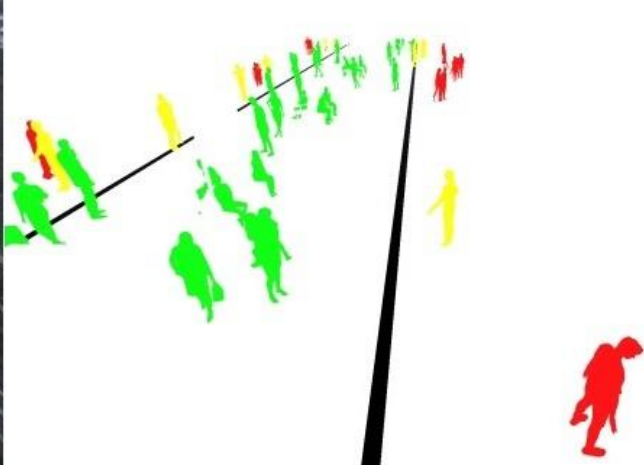
Unser erster Ansatz für die Klassifikation eines Bildes in das richtige Szenario, funktioniert durch ein Convolutional Neural Network (CNN), welches jedem Pixel eines Bildes eine der folgenden 5 Klassen zuweist:

- Weiß – Bahnhof mit fixen Objekten
- Schwarz – Sicherheitslinie
- Grün – Personen die keine Gefahr darstellen
- Gelb – Personen die eine potentielle Gefahr darstellen
- Rot – Personen die eine absolute Gefahr darstellen

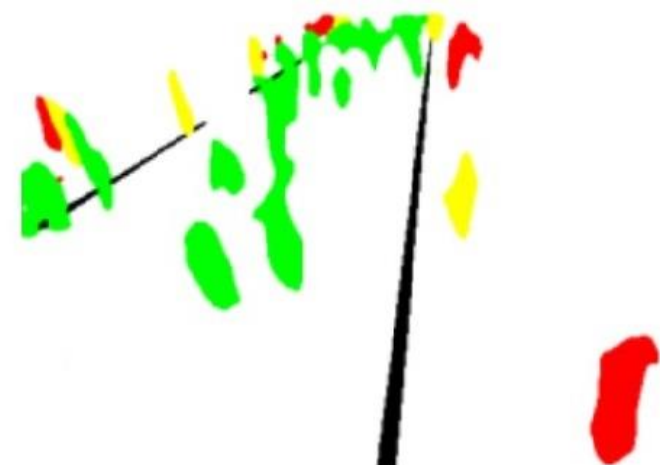
Das CNN wurde auf 50.000 Bildern mit insgesamt 1.000 verschiedenen Bahnhöfen aus unserer Simulation trainiert und anschließend auf 10.000 weiteren Bildern validiert. Im folgenden sind einige Beispielausgaben abgebildet.



Original



Ground Truth



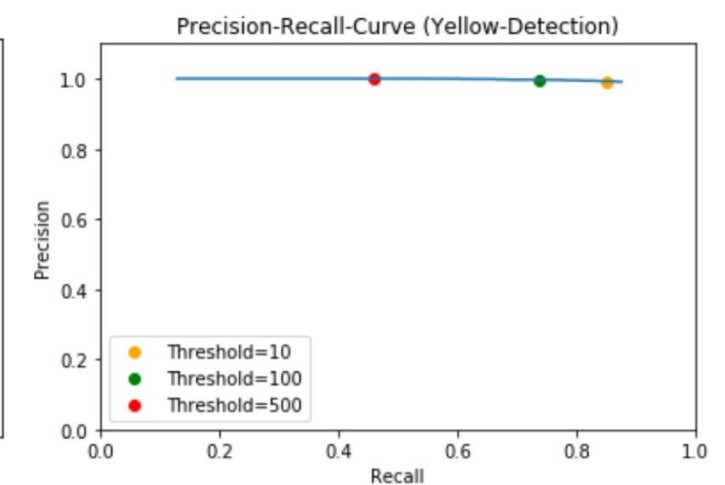
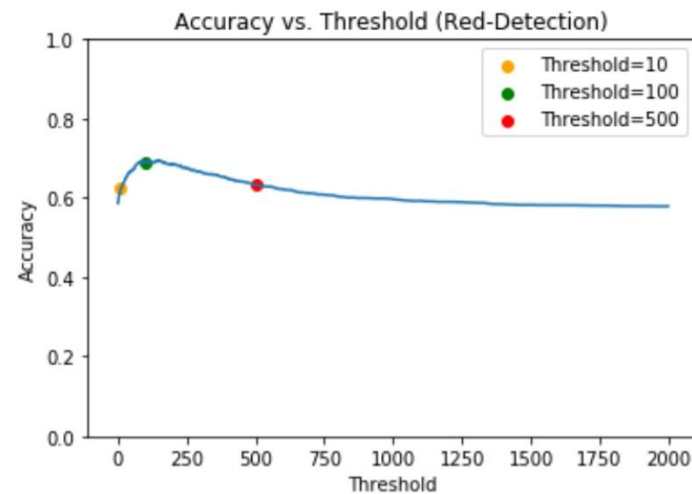
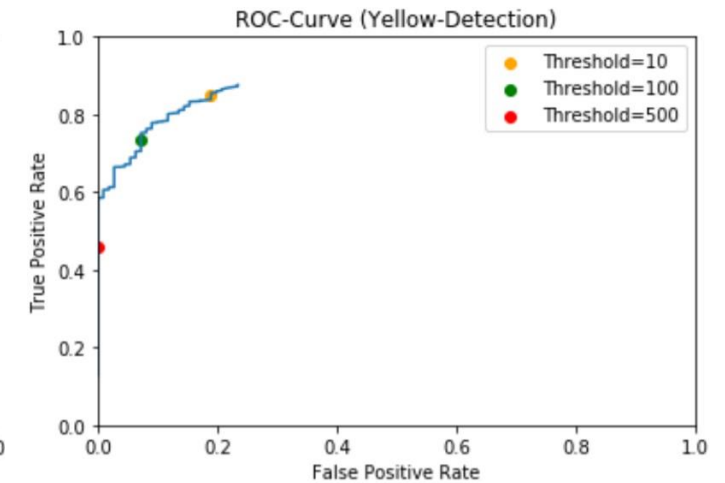
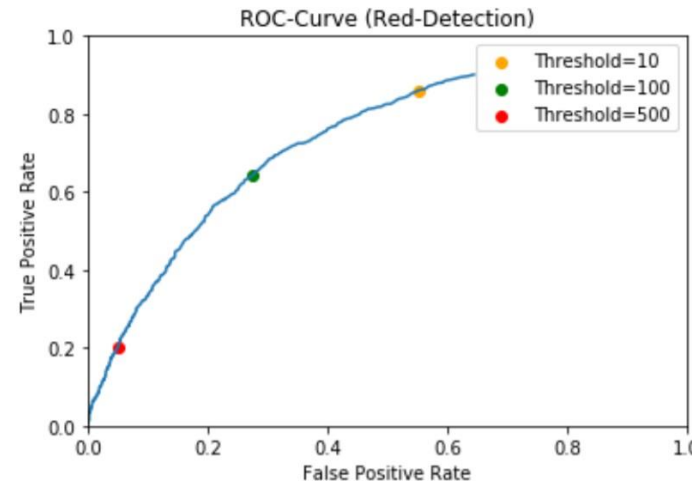
Prediction

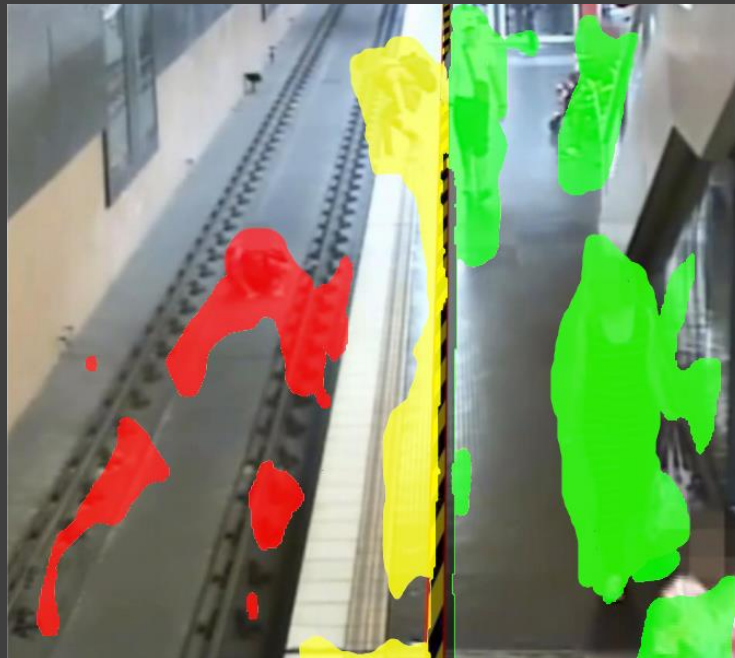
Erster Prototyp – Counting Algorithmus

Die Vorhersage des CNNs wird anschließend von einem von uns entwickeltem Counting Algorithmus verarbeitet und auf Gefahr untersucht. Der Algorithmus entscheidet anhand eines Schwellenwertes (Threshold) ob eines der folgenden Ereignisse eingetreten ist:

- Rot (es besteht absolute Gefahr)
- Gelb (es besteht potentielle Gefahr)

Das Ergebnis der Validierung unserer ersten Version zeigt, dass unser Verfahren funktioniert.





Erster Prototyp – Test auf echten Daten

Unser Modell wurde auf echten Daten einer Kölner U-Bahnstation getestet. Dafür haben wir ein Video vom “Kölner Schubser” verwendet, in dem eine Person von einer anderen auf das Gleis gestoßen wird. Das System erkennt die Personen in der Szene und klassifiziert sie in die richtige Klasse. Allerdings erhalten wir schlechte Objektgrenzen und einige Vorhersagen stimmen nicht.

Durch die ungünstigen Lichtverhältnisse ergeben sich unscharfe Objektgrenzen. In einer späteren Simulation werden wir unser System um realistische Lichtverhältnisse und Reflektionen erweitern.



Ausblick

Um die Trainings- und Validierungsdaten zu verbessern, erweitern wir die Simulation um weitere Objekte und Personen. Zudem fügen wir realistische Lichtverhältnisse und Reflektionen in die Simulation ein. Um die Trainingsdaten noch realistischer zu gestalten, nutzen wir zur Weiterverarbeitung Generative Adversarial Networks (GANs).

Für die Segmentierung des Bildes ziehen wir zwei neue Ansätze in Betracht. In beiden Ansätzen nutzen wir aus, dass die Kameras statisch sind und immer denselben Bahnhofsausschnitt zeigen. Im ersten Ansatz soll ein CNN ausschließlich die Sicherheitslinie und die Kante vom Bahnsteig erkennen. Die Personen im Bild werden durch Background Subtraction mit einem statistischen Ansatz zur Berücksichtigung veränderbarer Lichtverhältnisse vom Hintergrund getrennt. Beide Ausgaben werden anschließend algorithmisch zur gesuchten Segmentierung verarbeitet.

Im zweiten Ansatz geben wir in das CNN zusätzlich zur aktuellen Bahnhofssituation ein Backgroundbild ein. Somit hätten wird die Aufgabe ähnlich wie im ersten Verfahren gelöst, mit dem Unterschied, dass in diesem Fall das CNN die Backgroundsubtraction mit Berücksichtigung der Veränderbaren Lichtverhältnisse selbst lernt.

Als zusätzliche Option für die Erkennung der Menschen am Bahnsteig ziehen wir Object Detecion anstelle von Segmentierung in Betracht.

Unsere Anforderung an unser System im Bezug auf die Performance ist es eine True-Positive Rate von 95% und eine False-Negative Rate von unter 1% zu erreichen.